



HAL
open science

Genomic profiling of mycosis fungoides identifies patients at high risk of disease progression

Léa Fléchon, Inès Arib, Ankit K Dutta, Lama Hasan Bou Issa, Romanos Sklavenitis-Pistofidis, Rémi Tilmont, Chip Stewart, Romain Dubois, Stéphanie Poulain, Marie-Christine Copin, et al.

► To cite this version:

Léa Fléchon, Inès Arib, Ankit K Dutta, Lama Hasan Bou Issa, Romanos Sklavenitis-Pistofidis, et al.. Genomic profiling of mycosis fungoides identifies patients at high risk of disease progression. *Blood Advances*, 2024, 8 (12), pp.3109 - 3119. 10.1182/bloodadvances.2023012125 . inserm-04782305

HAL Id: inserm-04782305

<https://inserm.hal.science/inserm-04782305v1>

Submitted on 14 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Genomic profiling of mycosis fungoides identifies patients at high risk of disease progression

Léa Fléchon,^{1,*} Inès Arib,^{2,*} Ankit K. Dutta,^{3-5,*} Lama Hasan Bou Issa,¹ Romanos Sklavenitis-Pistofidis,³⁻⁵ Rémi Tilmont,² Chip Stewart,⁵ Romain Dubois,⁶ Stéphanie Poulain,^{1,7} Marie-Christine Copin,⁸ Sahir Javed,⁹ Morgane Nudel,² Doriane Cavalieri,² Guillaume Escure,² Nicolas Gower,² Paul Chauvet,² Nicolas Gazeau,² Cynthia Saade,² Marietou Binta Thiam,² Aïcha Ouelkite-Oumouchal,¹ Silvia Gaggero,¹ Émeline Cailliau,¹⁰ Sarah Faiz,¹¹ Olivier Carpentier,¹¹ Nicolas Duployez,^{1,7} Thierry Idziorek,¹ Laurent Mortier,^{11,12} Martin Figeac,¹³ Claude Preudhomme,^{1,7} Bruno Quesnel,^{1,2} Suman Mitra,¹ Franck Morschhauser,² Gad Getz,^{5,14,15} Irene M. Ghobrial,^{3,4,15} and Salomon Manier^{1,2}

¹Canther, ONCOLille, INSERM UMR-S1277, CNRS UMR9020, and ¹²OncoThai unit, INSERM UMR-S1189, Lille University, Lille, France; ²Department of Hematology, ⁶Department of Pathology, ⁷Department of Hematology, Biology and Pathology Center, ¹⁰Department of Biostatistics, and ¹¹Department of Pathology and Dermatology, Lille Hospital, Lille, France; ³Center for Prevention of Progression of Blood Cancers, Dana-Farber Cancer Institute, Boston, MA; ⁴Department of Medical Oncology, Harvard Medical School, Boston, MA; ⁵Cancer Program, Broad Institute of MIT and Harvard, Cambridge, MA; ⁸Department of Pathology, Angers University, Angers Hospital, INSERM, CRCI2NA, Angers, France; ⁹Department of Medical Oncology, Valenciennes Hospital, Valenciennes, France; ¹³Lille University, Lille Hospital, CNRS, INSERM, Institut Pasteur de Lille, US 41 – UAR 2014 - PLBS, Lille, France; ¹⁴Cancer Center and Department of Pathology, Massachusetts General Hospital, Boston, MA; and ¹⁵Harvard Medical School, Boston, MA

Key Points

- Genomic analysis of MF identifies alterations associated with high risk of progression and shorter overall survival.
- Clonal evolution of MF shows acquisition of *JUNB*, gain of 10p15.1 (*IL2RA/IL15RA*), or del12p13.1 (*CDKN1B*) at progression.

Mycosis fungoides (MF) is the most prevalent primary cutaneous T-cell lymphoma, with an indolent or aggressive course and poor survival. The pathogenesis of MF remains unclear, and prognostic factors in the early stages are not well established. Here, we characterized the most recurrent genomic alterations using whole-exome sequencing of 67 samples from 48 patients from Lille University Hospital (France), including 18 sequential samples drawn across stages of the malignancy. Genomic data were analyzed on the Broad Institute's Terra bioinformatics platform. We found that gain7q, gain10p15.1 (*IL2RA* and *IL15RA*), del10p11.22 (*ZEB1*), or mutations in *JUNB* and *TET2* are associated with high-risk disease stages. Furthermore, gain7q, gain10p15.1 (*IL2RA* and *IL15RA*), del10p11.22 (*ZEB1*), and del6q16.3 (*TNFAIP3*) are coupled with shorter survival. Del6q16.3 (*TNFAIP3*) was a risk factor for progression in patients at low risk. By analyzing the clonal heterogeneity and the clonal evolution of the cohort, we defined different phylogenetic pathways of the disease with acquisition of *JUNB*, gain10p15.1 (*IL2RA* and *IL15RA*), or del12p13.1 (*CDKN1B*) at progression. These results establish the genomics and clonality of MF and identify potential patients at risk of progression, independent of their clinical stage.

Introduction

Primary cutaneous T-cell lymphoma (CTCL) is a clinically heterogeneous group of incurable extranodal lymphomas that target skin-resident mature T cells. Mycosis fungoides (MF) is the most common CTCL, accounting for >50% of all cases.¹ It typically exhibits an indolent disease course with slow progression

Submitted 12 December 2023; accepted 17 February 2024; prepublished online on *Blood Advances* First Edition 21 March 2024. <https://doi.org/10.1182/bloodadvances.2023012125>.

*L.F., I.A., and A.K.D. contributed equally to this study.

The WES samples used in this study were deposited in the SRA database (<https://www.ncbi.nlm.nih.gov/sra>; accession number: PRJNA925900).

Deidentified individual participant data are included as supplements in the online version of this article. Other data are available upon request from the corresponding author, Salomon Manier (salomon.manier@chu-lille.fr).

The full-text version of this article contains a data supplement.

© 2024 by The American Society of Hematology. Licensed under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International \(CC BY-NC-ND 4.0\)](https://creativecommons.org/licenses/by-nc-nd/4.0/), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.

over several years, and patients present with a wide range of clinical symptoms and disease outcomes.² In the early stages, MF generally manifests as erythematous macules and plaques, which are nonspecific lesions that are difficult to diagnose. Patients may progress from low-risk (LR) clinical stages (tumor-node-metastasis-blood [TNMB] classification system IA to IIA) to high-risk (HR) stages (TNMB IIB to IVB), which involve tumors or generalized erythroderma.^{2,3} At the cellular level, ~20% of patients with HR show histological transformation with transformed MF cells, a feature associated with poor prognosis.⁴⁻⁶ Transformation is defined by the presence of >25% large cells (immunoblasts, large pleomorphic cells, or large anaplastic cells), which may or may not express CD30 within the infiltrate of the MF lesion.⁵ At the molecular level, MF is genetically heterogeneous, with no uniform single nucleotide variants (SNVs) and somatic copy number alterations (SCNAs). Previous microarray gene expression profiling studies of CTCLs, particularly MF, revealed deregulated expression of *TP53*, *PLCG1*,^{7,8} and *TNFR2*⁹ and deletions of the JAK-STAT signaling inhibitors *SOCS1* and *HNRNP1K*.¹⁰ Upon histological transformation, loss of chromosomal region 9p21.3, in which the tumor suppressors *CDKN2A* and *CDKN2B* reside, is commonly observed.¹⁰⁻¹³

Nevertheless, the molecular underpinnings of disease progression in patients with MF have not been extensively studied; thus, clinicians rely on suboptimal clinical variables for risk stratification. To identify the molecular drivers of disease progression and aggressiveness, which may help improve clinical practice, we collected samples from patients with MF at different stages of the disease and performed deep whole-exome sequencing (WES) for the detection of somatic events. Specifically, we sequenced DNA from 67 skin samples obtained from 48 patients, including 18 sequential samples from patients who exhibited progression from LR to HR stages. This approach allowed for us to characterize the genomic landscape of MF across the stages of progression, identify putative driver genes contributing to the progression of patients with LR, and study the clonal evolution of tumor cells during MF progression. Overall, this study provides new insights into the genetic risk factors of disease progression in patients with MF, which may help improve clinical prognostication models and lead to the discovery of novel therapeutic approaches for patients with MF. Although these findings identify patients with a HR of progression, further validation in independent cohorts is needed to confirm their clinical utility.

Methods

Patient samples

We studied 67 tumor samples from a cohort of 48 patients diagnosed with MF at the Lille University Hospital between 2003 and 2017. All cases were reviewed by local specialists from the French Study Group on Cutaneous Lymphomas Network. Immunohistochemistry was performed using the following panel: CD20, PAX5, CD2, CD3, CD4, CD5, CD7, CD8, CD30, PD1, and Ki-67. Another immunohistochemistry analysis of CD25 was performed in patient samples with or without gain10p15.1. The diagnosis of MF was confirmed by polymerase chain reaction to detect clonal recombination of the T-cell receptor gene. We defined 2 prognostic groups: the LR group with TNMB stages IA, IB, and IIA and the HR group with either high TNMB stage (IIB, III, and IV) or transformed histology. Samples were extracted from 67 frozen skin

biopsy samples stored in the Biology-Pathology-Genetics unit's tumor bank (certification NF 96900-2014/65453-1). After macrodissection of the tumors, the percentage of tumor cells was visually estimated by microscopic observation. In total, the cohort included 67 tumor samples from 48 patients. For 13 of these patients, sequential samples were collected at different stages of disease progression. This study was approved by the Institutional Review Board of Lille University Hospital and Nord Ouest IV (protocol number: ECH18/03) in accordance with the Declaration of Helsinki, and all patients provided written informed consent.

DNA quality control and WES

The amount of extracted DNA was determined by spectrophotometry using the "Quant-iT Picogreen dsDNA Assay" kit (Thermo Fisher Scientific). For WES, 100 ng of genomic DNA was fragmented on the Covaris ultrasonicator to target a base pair peak of ~150 to 200 bp. DNA libraries were prepared using the Agilent SureSelect XT low input kit with target region capture using Agilent SureSelect Human All Exon V7 (Agilent Technologies) and the Illumina Dual-Index Adapter primer kit (Illumina). Libraries were generated in an automated manner using the Bravo NGS liquid handling robot in accordance with the supplier's recommendations. The final libraries were assessed using the Agilent High Sensitivity DNA Analysis kit on the Bioanalyzer 2100 and quantified by quantitative polymerase chain reaction using the Kapa Library Quantification kit (Roche). Libraries were pooled and sequenced using 2 × 100 bp paired-end reads on an Illumina NovaSeq 6000 platform on a S4 flowcell.

Alignment and quality control

Data were analyzed on the Terra computing platform (Broad Institute [BI], Cambridge, MA), which aggregates bioinformatics tools for genomic data analysis. The quality of the raw sequencing output in fastq format was obtained using FastQC¹⁴ software and visualized at the whole-cohort level with MultiQC.¹⁵ Illumina primers, very small reads (< 30 bp), and poor quality reads were removed with CutAdapt.¹⁶ Sequences were aligned with BWA-MEM,¹⁷ and alignment quality was estimated using Picard software suite tools. Postalignment cleanup consisted of the removal of duplicate reads with MarkDuplicate¹⁸ (Picard, BI), local realignment around insertion-deletions (InDels) with IndelRealigner, and base quality recalibration with BaseRecalibrator and ApplyBQSR.¹⁹ During sample preparation in the laboratory, there were risks of contamination or inversion, which were checked with CalculateContamination²⁰ and CrossCheckLaneFingerprints,¹⁸ respectively, and no errors were identified. Finally, the potential oxidation of guanine to 8-oxoguanine artifacts that can occur during the preparation of genomic libraries under the combined effect of heat, DNA cutting, and the introduction of metal contaminants were removed with CollectOxoGMetrics.¹⁸ Overall, tumor samples meeting all quality control cutoffs had an average coverage of 99.72% on the GRCh37 assembly, with a mean target depth of coverage of 231.85x.

Copy number analysis from WES data

SCNAs and genome-wide allelic variations were detected using ModelSegments²⁰ software. To summarize, the software will use the panel of normal (PoN) of 13 samples to detect copy and allele number variations in each tumor sample and model several

segments. The software was run in Tumor-Only mode for the entire set of samples and run again in “Matched-Normal” mode for the 3 tumor samples that had an associated normal sample. Frequent large and focal SCNAs in the cohort were highlighted by GIS-TIC2.0²¹ using a q-value threshold of 0.01.

Mutation calling of recurrently mutated genes

SNVs and InDels were detected using a Bioinformatics Cancer Genome Analysis (CGA) pipeline called “CGA WES Characterization Pipeline.” SNVs were detected by MuTect²² and InDels by Strelka²³ and Mutect2²⁴. The results of these 3 software programs were then filtered using MAFFPoNFilter²⁵. The latter uses 2 PoNs to segregate somatic from germ line variants: the 1 from our cohort (from 13 WES normal samples) and the controlled access PoN used in the routine analysis of the BI (8334 normal samples from The Cancer Genome Atlas Program database). Somatic variants (SNVs and InDels) were annotated for their oncogenic effect using Variant Effect Predictor²⁶ and Oncotator.²⁷ They were then validated using the MutationValidator tool, which establishes the minimum number of reads carrying the variant for it to be considered somatic. Genes more frequently mutated than chance were determined using MutSig2CV²⁵. Variants carried by genes known to be “fishy genes,” that is, known false-positive genes that are not plausible in the development of cancers, were manually removed according to the list established by Lawrence et al in 2013.²⁸

Panel of normal

In addition to the 3 normal samples from our cohort, 10 additional human blood samples were provided by the Research Blood Component (Watertown, MA). These 13 samples formed a PoN to filter out possible errors during the preparation or sequencing steps, as well as germ line genetic/genomic events. As a final criterion to filter potential germ line variants, we used bash scripts to count the occurrence of mutated variants in our cohort of PoN samples in the fastq format. First, we converted the maf file into a 2-motive list per variant, consisting of 11 nucleotides before and after the variant. The first motive represented the wildtype, corresponding to the reference variant, whereas the second one was the mutated form, representing the alternative variant. Subsequently, we used the “do_it.sh” bash script from the GitHub repository mafouille/HotCount to count the number of occurrences of both wildtype and mutated motives in the PoN. Variants with mutated patterns that appeared more than once in the PoN were excluded. This additional check not only reveals supplementary potential artifacts but also modulates the stringency of the filtering process, complementing the PoN in the pipeline.

Estimation of purity, ploidy, and CCF

ABSOLUTE²⁹ software determined the purity and ploidy of each sample from the SNVs and SCNAs (supplemental Table 10) and thus served us to define the cancer cell fraction (CCF) of each genomic alteration. For tumor samples without associated normal samples, we applied a tumor-only Germline Somatic Log odds filter, as described by Chapuy et al in 2018.³⁰ To summarize, for each variant, its CCF, sample ploidy and purity, and local copy number (CN) were used to calculate the probability that the allelic fraction of the variant was consistent with a modeled allelic fraction for either a germ line or somatic hypothetical event. Thus, the filter set a threshold for each sample to remove additional germ line

variants. To eliminate additional germ line events, we set the tumor-only threshold to -1 .

GnomAD filtering

We used the genome aggregation database gnomAD to exclude potential germ line mutations. The Human Genome Variation Society genomic identifier (HGVS_genomic_change column in maf file) was used in the VEP GRCH37 Ensembl database to obtain gnomAD allele frequency. A cutoff value of 0.0001 was applied.

Statistical analysis

Time-to-event end points were estimated using the Kaplan-Meier method. Differences in the survival curves were assessed using the log-rank test. The median follow-up was calculated using the reverse Kaplan-Meier method. Time to progression was measured from the date of diagnosis to the date of documented progression to HR. Cox proportional hazards modeling was performed to assess the impact of genetic alterations on the risk of disease progression. Violin plot conditions were compared using the nonparametric Wilcoxon test. Figures and statistical estimations were obtained using R v.3.6.3 and MATLAB. The maf files were analyzed using the R package “mafTools” v2.12.05 with oncoplot (waterfall plot, Figure 1A), maf_compare (forest plot, Figure 2A), and somaticInteractions (corplot, Figure 3B) functions. The maf_compare function performs a Fisher test on all genes between the LR and HR cohorts to detect differentially mutated SNVs/SCNAs. In the generated forest plot, events with a low frequency of less than ~20% (ie, 9 patients) were not analyzed (Figure 2A). Other figures were generated using the R packages v3.4.0 “survival” v3.4.0 and “survminer” v0.4.9. (Kaplan-Meier graphs, Figure 2), and “chris-miller/fishplot” v0.5.1 (Fishplots, Figure 4) and ggplot2.

This study was approved by the Institutional Review Board of Lille University Hospital and Nord Ouest IV (protocol number: ECH18/03) in accordance with the Declaration of Helsinki, and all patients provided written informed consent.

Results

Genomic landscape of MF

The median age in our cohort was 62.5 years (range, 19-94), with a 60% male predominance, which is consistent with the known distribution of MF in the population.² Patients were stratified based on the TNMB classification into LR (stage IA, IB, or IIA) and HR (stages IIB to IVB) (supplemental Figure 1A-D; supplemental Tables 1 and 7). We detected mutations by WES and analyzed 67 tumor samples from 48 patients, using a validated pipeline to filter germ line variants and artifacts from tumor-only samples³⁰ (Methods; supplemental Figures 1 and 2). Thirteen patients had sequential sampling (supplemental Figure 1C-D).

We found a median of 3.5 mutations per Mb, corresponding to a median of 135 SNVs or insertion-deletions (InDels) per sample. The most recurrently mutated genes comprise previously reported mutational drivers in TCL, and 51% of patients had a mutation in at least 1 of these drivers (Figure 1A). These included the T-cell differentiation transcription factor *JUNB* (p.A282V; Figure 1B) in 13% of the cases³¹⁻³⁴; the epigenetic factor *TET2* in 9% of the cases³⁵⁻³⁸; the component of the MAPK pathway *MAPK1* in 6% of the cases (Figure 1B); the transcription factor *FOXA1* involved in

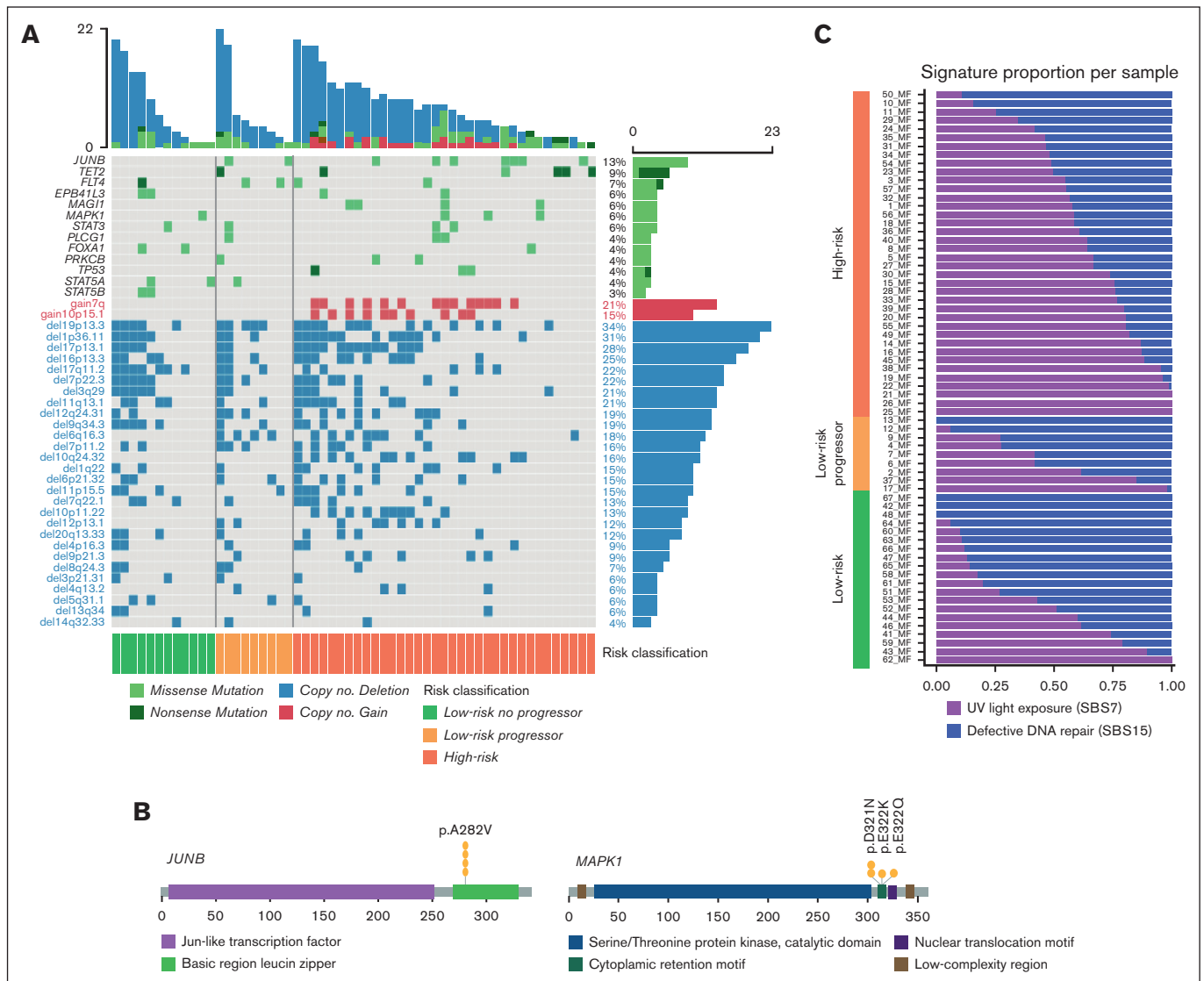


Figure 1. Landscape of most recurrent genomic alterations in MF. (A) Landscape of genomic alterations in 67 tumor samples of MF divided into LR and HR disease based on the TNMB stages. Alterations are divided into SNVs and SCNAs. (B) Driver oncogene maps of *JUNB* and *MAPK1*. (C) Relative enrichment of signature activities per samples divided into LR, LR progressors, and HR.

interferon signaling and immune response suppression^{37,39}; the tyrosine kinase receptor *FLT4*; the phospholipase *PLCG1* involved in NF- κ B and NFAT signaling^{7,40}; the JAK/STAT signaling pathway factors *STAT3*, *STAT5A*, and *STAT5B*; and *TP53* (Figure 1A; supplemental Figure 4; supplemental Table 2).

Next, we identified significantly recurrent SCNAs by using GIS-TIC2.0.²¹ Overall, SCNAs were the most common genomic alterations and were present in 84% of cases (Figure 1A). Specifically, we detected significantly recurrent alterations, including 8 arm-level and 29 focal copy number losses, as well as 2 arm-level and 1 focal copy gains (q-value ≤ 0.1 ; supplemental Figures 3A-B; supplemental Table 11). The frequencies of these SCNAs ranged from 6% to 57%, and the number of genes in the focal peaks varied from 1 (*MUC12* in del7q22.1) to 576 (del6p21.33) (supplemental Tables 12 and 13). In the focal CN gain *10p15.1* (26

genes; frequency of 13%) resides interleukin-2 receptor alpha and interleukin-15 receptor alpha (*IL2RA* and *IL15RA*) as well as the NF- κ B pathway protein kinase *PRKCB*. In particular, *IL2RA* and *IL15RA* play crucial roles in the phosphorylation of *STAT3* and *STAT5* in the JAK-STAT pathway, likely contributing to the pathogenesis of MF.⁴¹⁻⁴³ Among the tumor suppressor genes affected by the most frequent CN loss were *TMEM259* (19p13.3, frequency 57%), *TP53* (17p13.1, 41%), *SUZ12* and *NF1* (17q11.2, 33%), *NOTCH1* (9q34.3, 26%), *CARD11* (7p22.3, 24%), *ZEB1* (10p11.22, 17%), *TNFAIP3* (6q16.3, 15%), *CDKN2A* (9p21.3, 11%), and *CDKN1B* (12p12.2, 9%).

Mutational processes produce distinctive footprints called mutational signatures in the cancer genome that capture both DNA damages and repair mechanisms. We applied SignatureAnalyzer,⁴⁴ a tool that uses both the 3-bases mutational sequence

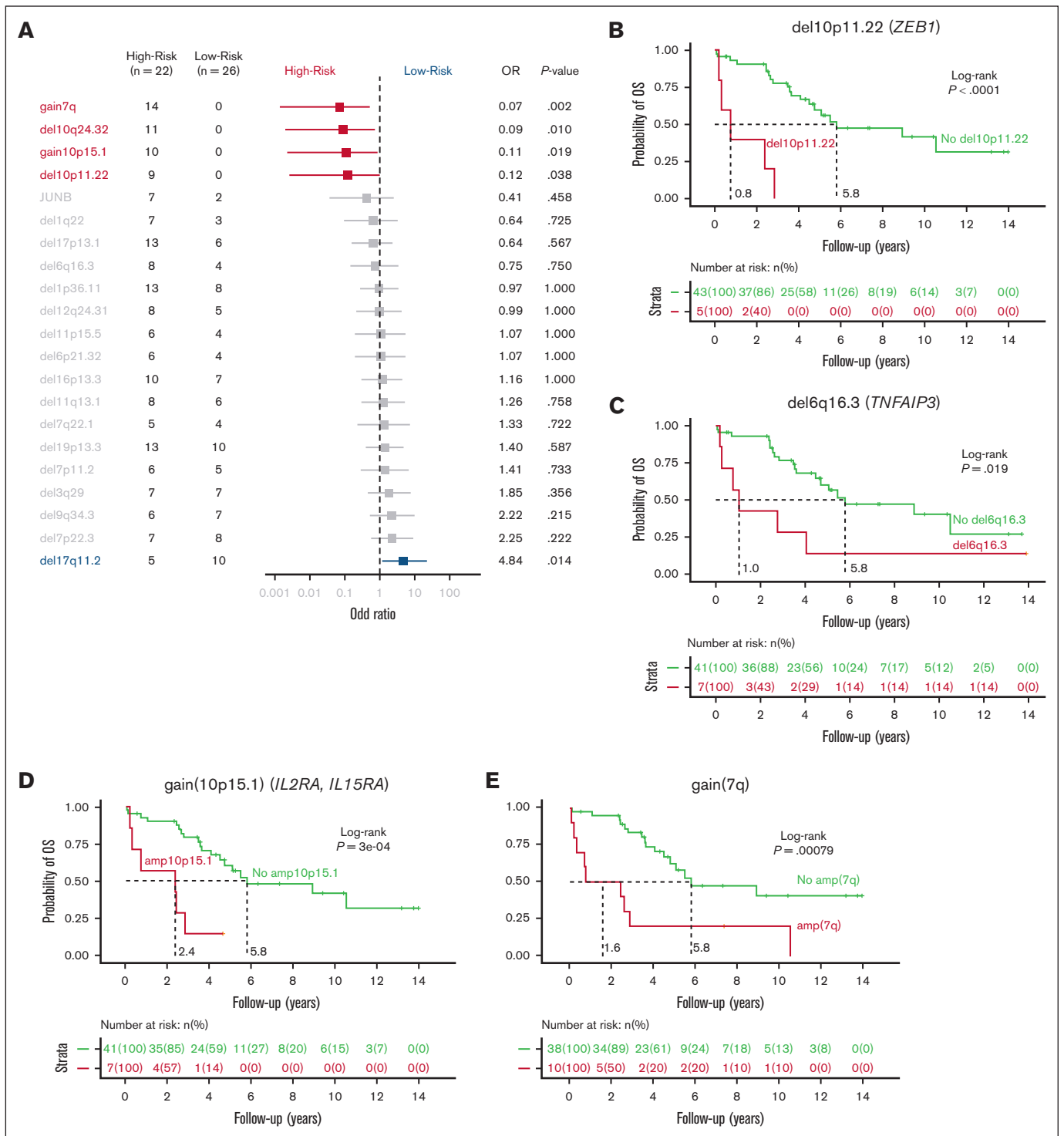


Figure 2. Correlation of genomic events to the disease stage. (A) Forest plot showing the association between individual genes alteration and clinical stage of MF divided into LR and HR, as depicted by OR. (B-E) Kaplan-Meier plots of individual genetic factors predictive of OS in univariate and multivariate models of 48 patients with a newly diagnosed MF: del10p11.22 (B); gain of 10p15.1 (C); gain of 7q (D); and del6q16.3 (E). (F) Kaplan-Meier curves for analysis of time to progression in patients with LR disease. P values were derived from log-rank test. (G) CD25 immunohistochemistry at diagnosis of MF skin biopsies in patients with gain of 10p15.1 (top panels, MF sample of patient 18 and patient 14 with presence of transformed MF cells) or without gain of 10p15.1 (bottom panels, MF sample of patient 23 and patient 16). Scale bars indicate 150 μ m.

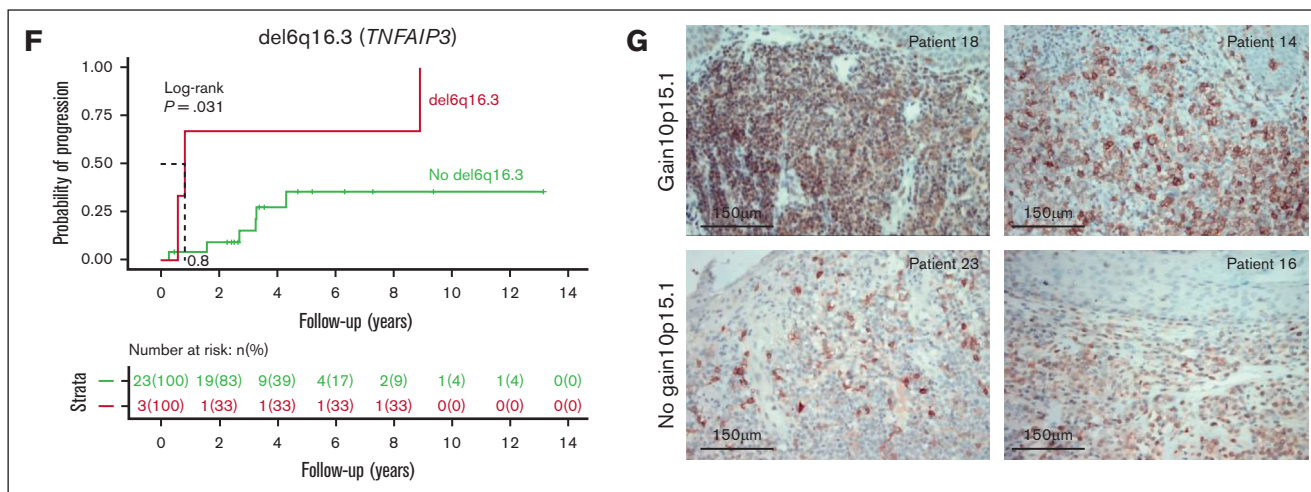


Figure 2 (continued)

context and the clustering of the mutation in the genome to define specific COSMIC (Catalogue of Somatic Mutations in Cancer database) single base substitution (SBS) signatures. Here, we detected 2 primary mutational signatures: the UV light exposure signature SBS7 and the defective DNA repair signature SBS15 (supplemental Figures 5A-B and E; supplemental Table 8). The age-related deamination signature, SBS1, was also frequently observed in our cohort but was not sufficiently prevalent to be separated from the UV signature. The enrichment of the SBS7 signature is consistent with previous studies on CTCL^{32,45,46} and was also significantly associated with HR, suggesting a role of UV radiation in the progression of MF. It is also possible that HR tumors are phylogenetically older and have more time to accumulate drivers and UV signature (Figure 1C). Additionally, there was an association between the total number of mutations and SBS7 signature (Wilcoxon test; $P = 5.735e^{-12}$). Signature SBS15 is 1 of the 7 mutational signatures related to defective DNA mismatch repair and microsatellite instability that are found in different cancer types. The contribution of SBS15 was similar in the LR and HR samples, suggesting a more founder event (supplemental Figure 5G). Next, we determined the relative contribution of SBS7 and SBS15 signatures to the mutational burden of driver genes (supplemental Figure 5E). Five genes were enriched with mutations associated with SBS15, including *JUNB*, *STAT3*, *MAML2*, *FLT4*, and *FOXA1*, whereas 9 genes were predominantly associated with SBS7, including *TET2*, *MAPK1*, *TP53*, *EPB41L3*, *MAGI1*, *PRKCB*, *HCK*, *L3MBTL4*, and *MECOM*.

Association of genetic features to disease stage and outcome

We further compared the genomic profiles of patients with LR disease with those of patients with HR disease (supplemental Tables 3-6 and 9). All CN gain7q (odds ratio [OR], 0.07; $P = .002$) and gain10p15.1 (*IL2RA* and *IL15RA*; OR, 0.11; $P = .019$) were found in patients at HR, as were del10q24.32 (*NFKB2*; OR, 0.09; $P = .010$) and del10p11.22 (*ZEB1*; OR, 0.12; $P = .038$; Figure 2A). Conversely, del17q11.2 (*SUZ12* and *NF1*) was significantly associated with LR (OR, 4.84; $P = .014$). We observed a higher mutation rate in HR samples, with a median of

3.95 mutations per Mb compared with 3.02 mutations per Mb in LR samples ($P = .019$) (supplemental Figure 5F). The increased mutation rate in HR was predominantly associated with the UV light exposure signature SBS7, suggesting either its role in disease progression or else reflecting time and addition of drivers (Figure 1C; supplemental Figure 5G).

Next, we assessed the impact of putative driver mutations at diagnosis on overall survival (OS) in patients with a median follow-up time of 4.3 years (range, 0.3-14). In total, 4 SCNAs were significantly associated with shorter OS, (Figure 2B-E): del10p11.22 (*ZEB1*; median OS [mOS] of 2.4 vs 8.9 years; $P = .00029$); gain10p15.1 (*IL2RA* and *IL15RA*; mOS of 2.4 vs 5.8 years; $P = .00029$); gain7q arm (mOS of 2.6 vs 8.9 years; $P = .011$); and del6q16.3 (*TNFAIP3*; mOS, 1.9 vs 5.8 years; $P = .021$). Furthermore, the presence of del10p11.22, gain10p15.1, and del6q16.3 remained a significant risk factor in a multivariate stepwise analysis, accounting for the patient's clinical stage (supplemental Figures 5C-D). Notably, no SNVs were shown to affect OS. In the LR group, 7 patients progressed, and 19 did not progress during follow-up. Patients with LR disease and del6q16.3 (*TNFAIP3*) had shorter time to progression (median, 0.8 years vs not reached; $P = .0031$; Figure 2F). This alteration is present in 15% of patients with LR at diagnosis and could serve as a prognostic marker in clinical practice for future publication with a validation cohort.

Regarding gain10p15.1, we observed a higher expression of CD25 by immunohistochemistry in samples from patients with gain10p15.1 than those of patients without gain10p15.1. This is consistent with a link between gain of *IL2RA* and the surface expression of CD25, suggesting a role of *IL2/IL2RA* signaling in the disease progression (Figure 2G).

The clonal architecture and phylogeny of MF

Next, we estimated the CCF for each putative driver and determined whether the alterations were clonal (≥ 0.9) or subclonal (< 0.9). We observed heterogeneity in the clonality of genomic alterations in patients with MF. Certain alterations were frequently clonal, such as del9p21.3 (*CDKN2A*), del10p11.22 (*ZEB1*), gain7q, and mutations in *PRKCB* or *STAT5B*. Other alterations

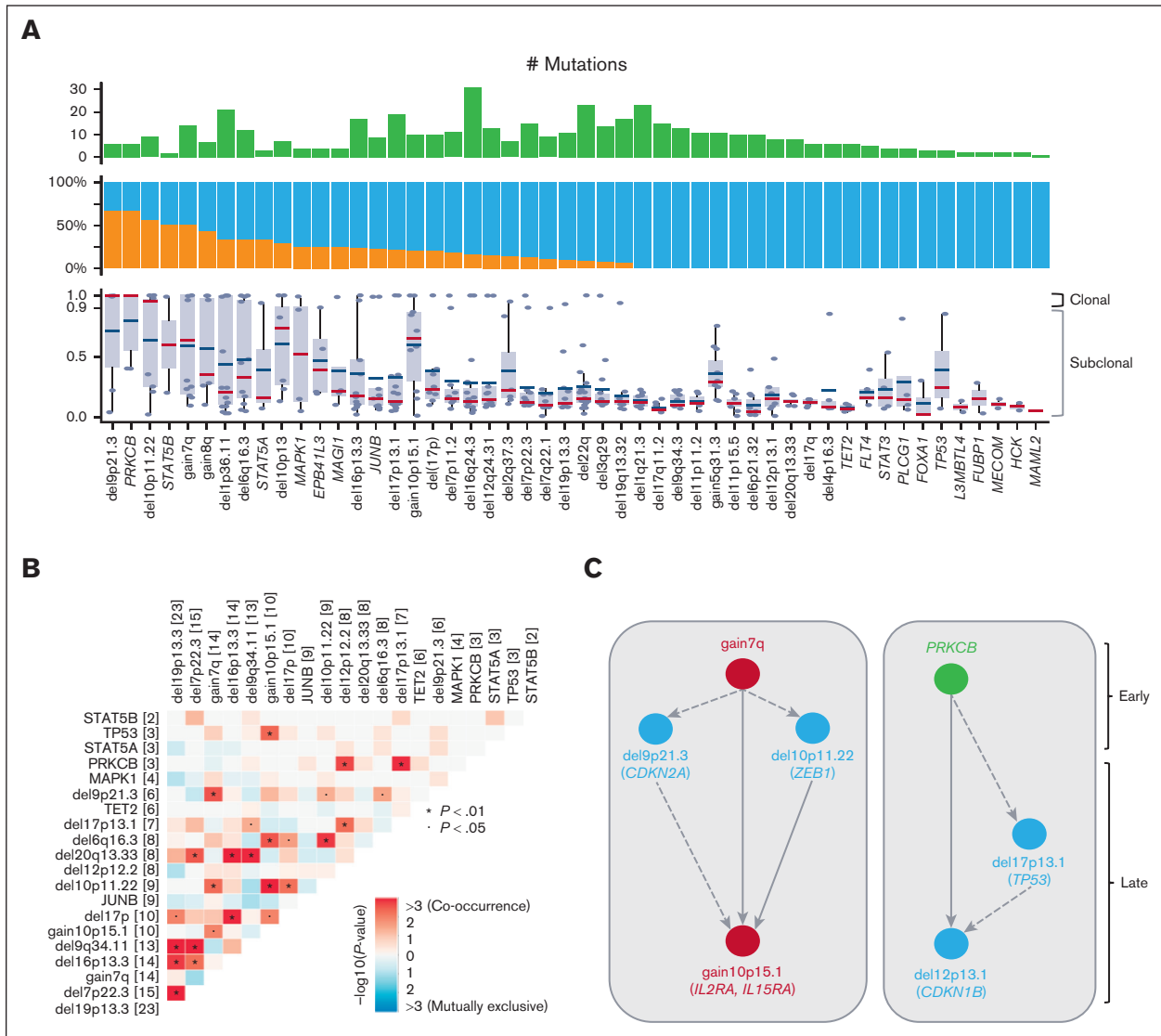


Figure 3. Clonal heterogeneity and inferred timing of genetic drivers. (A) Proportion in which recurrent drivers are found as clonal or subclonal across the 67 samples (top), along with the individual CCF values for each sample affected by a driver (bottom). Median CCF values are shown (bottom, bars represent the median and interquartile range for each driver). (B) Correlation matrix for the most recurrent genomic alterations for the 48 patients with MF, with Fisher exact test to detect such significant pair of mutations. (C) Timing of genomic alterations with early events at top and late events at bottom. Color indicates alteration types. Arrows between 2 alterations were drawn when 2 drivers were found in 1 sample with an excess of clonal to subclonal events. Dashed arrows indicate 1 clonal-subclonal pair and solid arrows indicate ≥ 2 clonal-subclonal pairs.

were more frequently subclonal, such as mutations in *TET2*, *TP53*, *PLCG1*, or *del17p*, which likely represent later events in the disease course (Figure 3A).

Subsequently, we analyzed the co-occurrence of driver genes and SCNAs. We found that *del9p21.3* (*CDKN2A*) significantly co-occurred with *gain7q* and *del10p11.22* (*ZEB1*). *Gain10p15.1* (*IL2RA* and *IL15RA*) was significantly cosegregated with *del6q16.3* (*TNFAIP3*) and *TP53* mutations. *Del16p13.3* (*CREBBP*) co-occurred with *del17p* and *del19p13.3* (*TMEM259*; Figure 3B). We applied a mutation-ordering method to samples with pairs of clonal and subclonal alterations.⁴⁷ Given that clonal mutations occur before subclonal events, we defined the timing of the main genetic alterations. We observed 2 general patterns

defining the phylogeny of the disease (Figure 3C): first, clonal *gain7q* that often co-occurs with clonal *del9p21.3* (*CDKN2A*) or *del10p11.22* (*ZEB1*), with further acquisition of *gain10p15.1* (*IL2RA* and *IL15RA*) at progression; second, the clonal mutation of the tyrosine kinase *PRKCB* followed by *del17p13.1* (*TP53*) and *del12p13.1* (*CDKN1B*). Of note, *gain7q* and *del10p11.22* (*ZEB1*) are often clonal and associated with an unfavorable outcome suggesting that the molecular path to aggressiveness is made at an early stage of the disease. *Gain10p15.1* (*IL2RA* and *IL15RA*), which is also associated with HR, represents a transforming molecular event in the disease course.

We further analyzed sequential samples from 7 patients who progressed from early stages (LR progressors) to advanced stages of

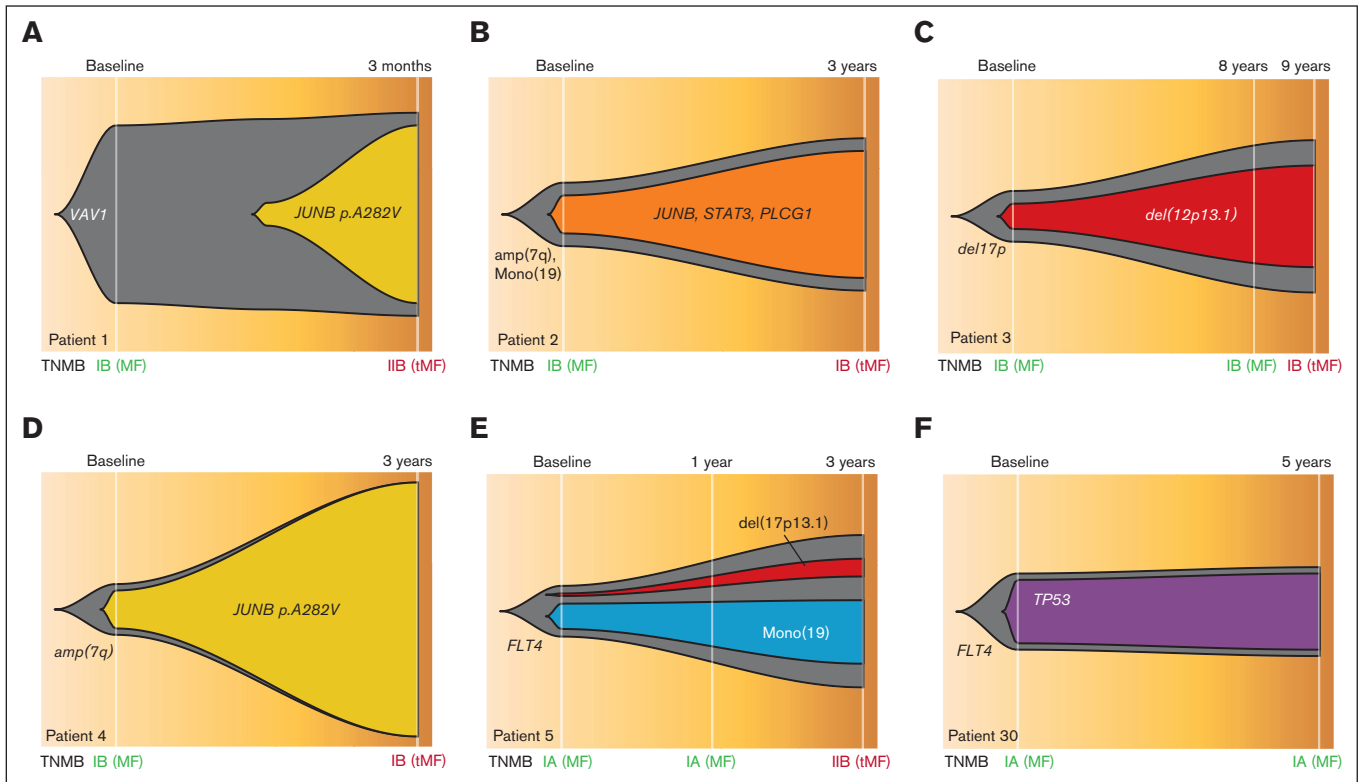


Figure 4. Clonal evolution of sequential samples. (A-E) Fish plots of 5 serial cases of MF with samples at LR stages who progressed to HR stages. (F) One serial case with both time points at LR stage because the patient did not progress. TNMB classification has been indicated (green for LR and dark red for HR samples). tMF, transformed MF.

the disease, with sampling intervals ranging from 3 months to 9 years. We observed evidence of clonal heterogeneity at diagnosis in all 7 cases, indicating that clonal branching had already occurred at the early stage of the disease. Moreover, we noticed cases in which subclones that were initially small at diagnosis exhibited substantial expansion at a later time point, indicating a strong driver potential for their mutations/SCNAs. Specifically, this was observed in 3 patients with small subclones harboring the *JUNB* mutation or *del12p13.1* (*CDKN1B*), which were selected for and became dominant at disease progression (Figure 4B-D). In these cases, all alterations detected in late-stage samples were already present at the baseline, pointing to a linear pattern of clonal evolution. In another instance, however, a *JUNB*-mutant subclone was first identified at the time of disease progression, suggesting that further branching had occurred and drove the progression from LR to HR disease (Figure 4A). In this case, the *JUNB* mutation was acquired on top of a clonal *VAV1* mutation, involved in T-cell receptor signaling, and has been reported in various T-cell malignancies (Figure 5). Clonal evolution was also observed during treatment, with the expansion of a resistant *del17p13.1* subclone in a patient receiving multiple lines of treatment within 3 years of disease progression (Figure 4E). Finally, in a patient who did not progress and whose 2 samples were sequenced at the LR stage but 5 years apart, no new alterations or increases in the CCF of baseline mutations were noticed at the later time point, implying that clonal heterogeneity at baseline may not be sufficient for disease progression (Figure 4F).

Discussion

In this study, we leveraged WES data to analyze tumor samples from 48 patients with newly diagnosed MF, together with sequential samples at later time points from 13 patients. We analyzed the most recurrent genomic alterations, as well as their clonal heterogeneity and clonal evolution, to temporally order these alterations and gain insight into the phylogeny of MF. Our results highlight the complexity of MF, with a median of 135 different genomic alterations per tumor. The most recurrent alterations in MF were consistent with previous observations in TCL in general^{9,32,36,38,48,49} (MF, Sézary syndrome, primary cutaneous CD30⁺ T-cell lymphoproliferative disorders, and primary cutaneous $\gamma\delta$ T-cell lymphoma), with alterations in the NF- κ B pathway (such as deletion *NFKB2* in 10q24.32, gain of *PRKCO* in 10p15.1, mutation of *PRCKB*, and deletion of *CARD11* in 7p22.3), JAK/STAT pathway (eg, gain of *IL2RA* and *IL15RA* in 10p15.1, and mutations in *STAT3*, *STAT5A* and *STAT5B*), MAPK pathway (mutations in *JUNB* and *MAPK1*), cell cycle pathway (such as deletion *CDKN1B* in 12p13.1 and *CDKN2A* in 9p21.3), and inhibition of apoptosis (deletion and mutations of *TP53* and deletion of *TNFAIP3* in 6p16.3; Figure 5).

We identified genomic alterations that are associated with HR stages of the disease: gain7q and gain10p15.1 (*IL2RA* and *IL15RA*), *del10q24.32* (*NFKB2*), and *del10p11.22* (*ZEB1*), suggesting their role in disease progression. Conversely, *del17q11.2* (*SUZ12* and *NF1*) was associated with the LR stages of the disease. The mutational signature SBS7, associated with UV light, was

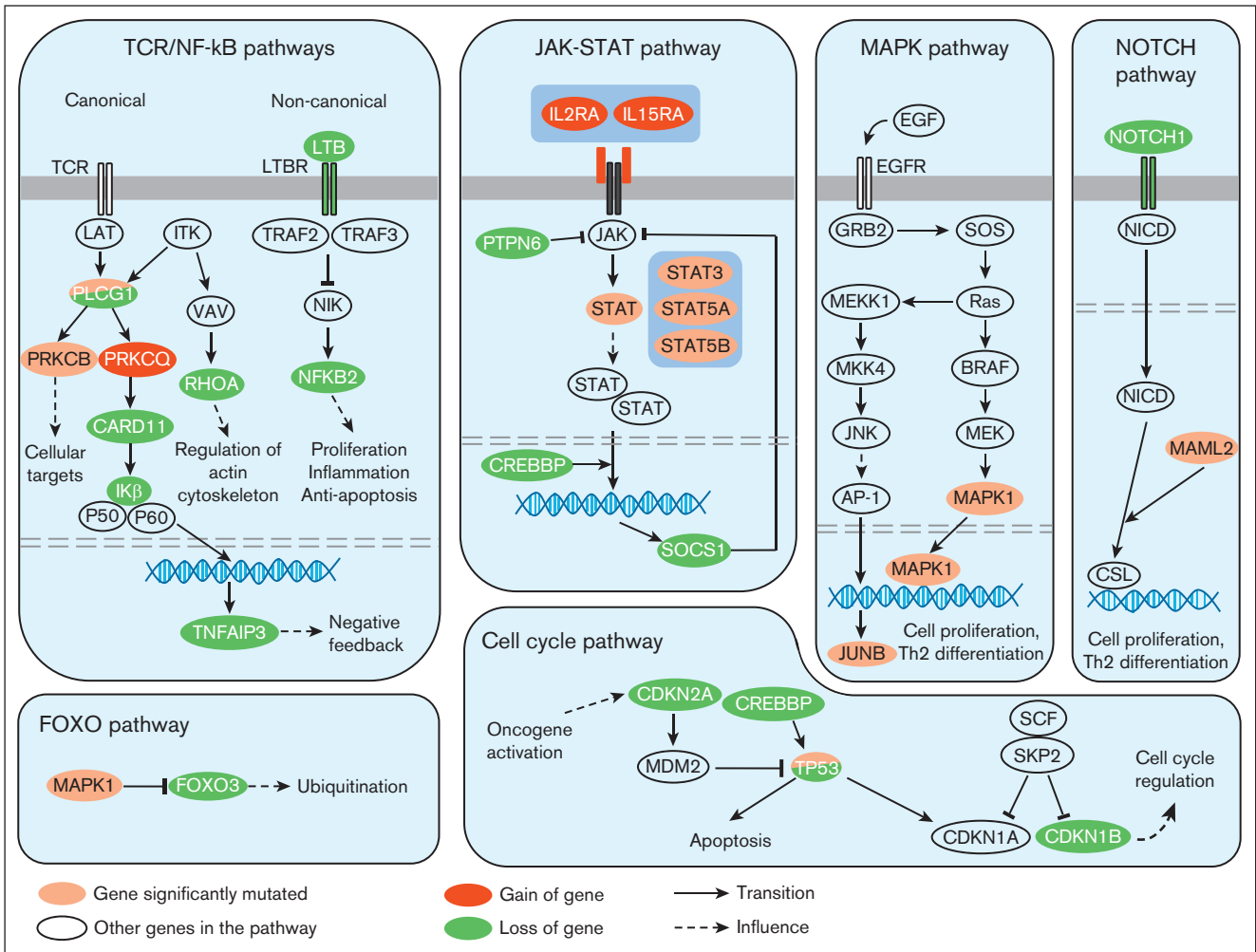


Figure 5. Recurrent genetic alterations affecting key signaling pathways involved in MF. In this diagram, frequently mutated genes with well-established roles in these signaling pathways have been depicted representing the proteins they encode. TCR, T-Cell Receptor.

also associated with HR stages and was enriched in the mutational process of genes such as *TET2*, *MAPK1*, *TP53*, and *PRKCB*. We also describe the different patterns of phylogeny of MF, with early events such as gain7q or mutations in *PRKCB* and further acquisition of alterations such as gain10p15.1 (*IL2RA* and *IL15RA*), del17p13.1 (*TP53*), or del12p13.1 (*CDKN1B*). This clonal evolution from the LR to HR stages was either linear or branched, with the selection of subclones with strong driving potential.

We also assessed the prognostic value of these genomic alterations during diagnosis. We found 4 SCNAs that were significantly associated with a shorter OS: del10p11.22 (*ZEB1*), gain10p15.1 (*IL2RA*, *IL15RA*), gain7q, and del6q16.3 (*TNFAIP3*). The identification of these HR genomic alterations can help in the clinical management of MF. Early focal lesions of MF are sometimes difficult to diagnose and histologically differentiate from eczematous or psoriasis dermatitis.⁵⁰ Specific genomic alterations associated with adverse outcomes can help to guide the diagnosis of MF. Our data illustrate the role of gain10p15.1 (*IL2RA* and *IL15RA*) in disease progression. Taken together, these genomic alterations represent potential early indicators that may be useful for MF prognostication but require validation in future studies. This can be therapeutically relevant, because

cytokine inhibitors such as bnz-1 are currently being tested in clinical trials.

In conclusion, sequencing a large cohort of patients with MF, including sequential samples, has allowed for us to understand the genomic complexity of MF, temporal ordering of genetic events, and biomarkers that are associated with HR and disease progression. We believe that introducing NGS evaluation at the time of MF diagnosis can improve the identification of patients at a HR of disease progression and their clinical management.

Acknowledgments

This study was supported by a grant from the GEFLUC Flandres Artois Foundation to (S. Manier). The authors acknowledge the Tumorotheque of CHU Lille, the Société Française d'Hématologie, and the I-SITE ULNE foundation for their assistance with this project.

Authorship

Contribution: L.F., I.A., and S. Manier contributed in conceptualization, methodology, and writing including original draft and

visualization; L.F., R.S.-P., C. Stewart and M.F. contributed in software; L.F., I.A., A.K.D., R.S.-P., C. Stewart, E.C., M.F., S. Mitra, G.G., I.M.G., and S. Manier contributed in formal analysis; L.F., I.A., A.K.D., L.H.B.I., and S. Manier contributed in investigation; I.A., R.D., S.P., M.-C.C., S.J., M.N., A.O.-O., S.F., O.C., L.M., M.F., T.I., F.M., G.G., I.M.G., and S. Manier contributed in resources; R.D. contributed in validation; L.F. and S. Manier contributed in data curation; all authors contributed in reviewing and editing; G.G., I.M.G., and S. Manier supervised the research; and S. Manier contributed in funding acquisition.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: L.F., [0000-0003-4140-475X](#); I.A., [0009-0000-7062-9814](#); A.K.D., [0000-0003-2535-3729](#); L.H.B.I., [0000-0002-2561-9980](#); R.S.-P., [0000-0002-8025-9850](#); R.T., [0000-0003-](#)

[1739-2162](#); C.S., [0000-0003-2245-9552](#); R.D., [0000-0001-6990-7354](#); S.P., [0000-0002-1792-7270](#); M.-C.C., [0000-0002-4729-7054](#); S.J., [0000-0002-1649-9577](#); M.N., [0000-0001-9100-2316](#); D.C., [0000-0002-4836-7237](#); G.E., [0000-0002-3151-9441](#); P.C., [0000-0001-6326-9283](#); N.G., [0000-0003-3449-6676](#); M.B.T., [0000-0002-1387-0510](#); A.O.-O., [0000-0002-7048-5555](#); S.G., [0000-0002-2685-1108](#); É.C., [0000-0001-5343-6471](#); O.C., [0000-0002-7788-6132](#); N.D., [0000-0002-3927-1022](#); T.I., [0000-0001-5893-9131](#); L.M., [0000-0002-2644-1790](#); M.F., [0000-0001-8523-3708](#); C.P., [0000-0002-1267-9546](#); B.Q., [0000-0002-6563-2709](#); S. Mitra, [0000-0002-3426-371X](#); F.M., [0000-0002-3714-9824](#); G.G., [0000-0002-0936-0753](#); I.M.G., [0000-0001-7361-3092](#); S. Manier, [0000-0001-7653-711X](#).

Correspondence: Salomon Manier, CHU Lille, Hospital Huriez, Rue Michel Polonovski, Lille 59000, France; email: salomon.manier@inserm.fr and salomon.manier@chu-lille.fr.

References

1. Willemze R. Cutaneous T-cell lymphoma: epidemiology, etiology, and classification. *Leuk Lymphoma*. 2003;44(Suppl 3):S49-S54.
2. Willemze R, Jaffe ES, Diaz-Perez JL, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768-3785.
3. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sézary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood*. 2007;110(6):1713-1722.
4. Diamandidou E, Colome-Grimmer M, Fayad L, Ducic M, Kurzrock R. Transformation of mycosis fungoides/sezary syndrome: clinical characteristics and prognosis. *Blood*. 1998;92(4):1150-1159.
5. Vergier B, de Muret A, Beylot-Barry M, et al. Transformation of mycosis fungoides: clinicopathological and prognostic features of 45 cases. *Blood*. 2000;95(7):2212-2218.
6. Benner MF, Jansen PM, Vermeer MH, Willemze R. Prognostic factors in transformed mycosis fungoides: a retrospective analysis of 100 cases. *Blood*. 2012;119(7):1643-1649.
7. Vaqué JP, Gómez-López G, Monsálvez V, et al. PLCG1 mutations in cutaneous T-cell lymphomas. *Blood*. 2014;123(13):2034-2043.
8. Caumont C, Gros A, Boucher C, et al. PLCG1 gene mutations are uncommon in cutaneous T-cell lymphomas. *J Invest Dermatol*. 2015;135(9):2334-2337.
9. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. *Nat Genet*. 2015;47(9):1056-1060.
10. Bastidas Torres AN, Cats D, Mei H, et al. Genomic analysis reveals recurrent deletion of JAK-STAT signaling inhibitors HNRNPK and SOCS1 in mycosis fungoides. *Genes Chromosomes Cancer*. 2018;57(12):653-664.
11. Böni R, Xin H, Kamarashev J, et al. Allelic deletion at 9p21-22 in primary cutaneous CD30+ large cell lymphoma. *J Invest Dermatol*. 2000;115(6):1104-1107.
12. Laharanne E, Chevret E, Idrissi Y, et al. CDKN2A-CDKN2B deletion defines an aggressive subset of cutaneous T-cell lymphoma. *Mod Pathol*. 2010;23(4):547-558.
13. Nicolae-Cristea A r, Benner M f, Zoutman W h, et al. Diagnostic and prognostic significance of CDKN2A/CDKN2B deletions in patients with transformed mycosis fungoides and primary cutaneous CD30-positive lymphoproliferative disease. *Br J Dermatol*. 2015;172(3):784-788.
14. Andrews S, Fast QC. A quality control tool for high throughput sequence data. 2010. Accessed 22 November 2022. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
15. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32(19):3047-3048.
16. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17(1):10-12.
17. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*. 2013.
18. Picard Tools (Broad Institute, Cambridge, MA). Version 2.27.5 (October 7, 2022). Accessed 12 February 2023. <http://broadinstitute.github.io/picard/>
19. Van der Auwera GA, Carneiro MO, Hartl C, et al. From fastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43(1110):11.10.1-11.10.33.
20. Van der Auwera GA, O'Connor BD, Brian D. *Genomics in the cloud : using Docker, GATK, and WDL in Terra*. O'Reilly Media; 2020.

21. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhir R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 2011;12(4):R41.
22. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol.* 2013;31(3):213-219.
23. Saunders CT, Wong WSW, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics.* 2012;28(14):1811-1817.
24. Benjamin D, Sato T, Cibulskis K, et al. Calling somatic SNVs and indels with mutect2. *bioRxiv.* Preprint posted online 2 December 2019. <https://doi.org/10.1101/861054>
25. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature.* 2014;505(7484):495-501.
26. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome Biol.* 2016;17(1):122.
27. Ramos AH, Lichtenstein L, Gupta M, et al. Oncotator: cancer variant annotation tool. *Hum Mutat.* 2015;36(4):E2423-E2429.
28. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013;499(7457):214-218.
29. Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol.* 2012;30(5):413-421.
30. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med.* 2018;24(5):679-690.
31. Luchtel RA, Zimmermann MT, Hu G, et al. Recurrent MSCE116K mutations in ALK-negative anaplastic large cell lymphoma. *Blood.* 2019;133(26):2776-2789.
32. Park J, Daniels J, Wartewig T, et al. Integrated genomic analyses of cutaneous T-cell lymphomas reveal the molecular bases for disease heterogeneity. *Blood.* 2021;138(14):1225-1236.
33. Mao X, Orchard G, Mitchell TJ, et al. A genomic and expression study of AP-1 in primary cutaneous T-cell lymphoma: evidence for dysregulated expression of JUNB and JUND in MF and SS. *J Cutan Pathol.* 2008;35(10):899-910.
34. Mao X, Orchard G, Lillington DM, Russell-Jones R, Young BD, Whittaker SJ. Amplification and overexpression of JUNB is associated with primary cutaneous T-cell lymphomas. *Blood.* 2003;101(4):1513-1519.
35. Lio C-WJ, Yuita H, Rao A. Dysregulation of the TET family of epigenetic regulators in lymphoid and myeloid malignancies. *Blood.* 2019;134(18):1487-1497.
36. Iyer A, Hennessey D, O'Keefe S, et al. Independent evolution of cutaneous lymphoma subclones in different microenvironments of the skin. *Sci Rep.* 2020;10(1):15483.
37. Iyer A, Hennessey D, O'Keefe S, et al. Branched evolution and genomic intratumor heterogeneity in the pathogenesis of cutaneous T-cell lymphoma. *Blood Adv.* 2020;4(11):2489-2500.
38. da Silva Almeida AC, Abate F, Khiabani H, et al. The mutational landscape of cutaneous T-cell lymphoma and Sézary syndrome. *Nat Genet.* 2015;47(12):1465-1470.
39. He Y, Wang L, Wei T, et al. FOXA1 overexpression suppresses interferon signaling and immune response in cancer. *J Clin Invest.* 2021;131(14):e147025.
40. Patel VM, Flanagan CE, Martins M, et al. Frequent and persistent PLCG1 mutations in sézary cells directly enhance PLCγ1 activity and stimulate NFκB, AP-1, and NFAT signaling. *J Invest Dermatol.* 2020;140(2):380-389.e4.
41. Zhang Q, Nowak I, Vonderheid EC, et al. Activation of Jak/STAT proteins involved in signal transduction pathway mediated by receptor for interleukin 2 in malignant T lymphocytes derived from cutaneous anaplastic large T-cell lymphoma and Sezary syndrome. *Proc Natl Acad Sci U S A.* 1996;93(17):9148-9153.
42. Eriksen KW, Kalltoft K, Mikkelsen G, et al. Constitutive STAT3-activation in Sezary syndrome: tyrphostin AG490 inhibits STAT3-activation, interleukin-2 receptor expression and growth of leukemic Sezary cells. *Leukemia.* 2001;15(5):787-793.
43. Qin J-Z, Kamarashev J, Zhang C-L, Dummer R, Burg G, Döbbeling U. Constitutive and interleukin-7- and interleukin-15-stimulated DNA binding of STAT and novel factors in cutaneous T cell lymphoma cells. *J Invest Dermatol.* 2001;117(3):583-589.
44. Kasar S, Kim J, Impropio R, et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat Commun.* 2015;6(1):8866.
45. Jones CL, Degasperi A, Grandi V, et al. Spectrum of mutational signatures in T-cell lymphoma reveals a key role for UV radiation in cutaneous T-cell lymphoma. *Sci Rep.* 2021;11(1):3962.
46. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood.* 2015;126(4):508-519.
47. Landau DA, Tausch E, Taylor-Weiner AN, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature.* 2015;526(7574):525-530.
48. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nat Genet.* 2015;47(9):1011-1019.
49. Argyropoulos KV, Pulitzer M, Maura F, et al. Targeted genomic analysis of cutaneous T cell lymphomas identifies a subset with aggressive clinicopathological features. *Blood Cancer J.* 2020;10(11):116.
50. Hristov AC, Tejasvi T, Wilcox RA. Mycosis fungoides and Sézary syndrome: 2019 update on diagnosis, risk-stratification, and management. *Am J Hematol.* 2019;94(9):1027-1041.